

УДК 681.3.06

ОПТИМИЗАЦИЯ ФУНКЦИОНИРОВАНИЯ МНОГОКАНАЛЬНЫХ РАСПРЕДЕЛЁННЫХ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ ДЛЯ ТРУДНОДОСТУПНЫХ ОБЪЕКТОВ

М.А. Сонькин, А.А. Шамин

Томский политехнический университет

Томский научный центр СО РАН

E-mail: salex@cc.tpu.edu.ru

Выбраны критерии оптимизации передачи данных в системах сбора гидрометеорологической информации. Предложены способы оценки времени и стоимости доставки сообщений по заданному каналу связи на основании статистических данных с учётом выбранного способа тарификации связи. Выработаны алгоритмы оптимизации передачи данных: по времени доставки, стоимости доставки и надёжности.

Ключевые слова:

Пакетная передача данных, оптимизация систем передачи данных, прикладная математика.

Особое место среди информационных территориально распределённых систем [1] занимают системы передачи данных для труднодоступных объектов — например, системы сбора гидрометеорологической информации. Особенностью таких систем передачи данных является, прежде всего, использование беспроводных каналов связи — радиоканала, спутниковых и сотовых каналов.

В связи с этими требованиями особое значение приобретает задача оптимизации системы связи по таким параметрам, как время, стоимость или надёжность доставки сообщения.

Существующие подходы к оптимизации на основе автоматического выбора каналов связи ориентированы на совокупность однородных каналов. Но, в рассматриваемых системах передачи гидрометеорологических данных используется, как правило, совокупность разнородных каналов связи, что делает существующие подходы неприменимыми.

В данной статье рассматривается возможность автоматического выбора канала связи в системе передачи данных, предназначенной для сбора гидрометеорологической информации на основе математической модели.

Рассматриваемая система [2] (рис. 1) представляет собой иерархическую древовидную структуру, элементами которой являются метеостанции (МС), центры сбора данных (ЦСД), управления ги-

дрометеослужбы (УГМС), центры коммутации сообщений (ЦКС) и каналы связи.

На верхнем уровне иерархии находится ЦКС Росгидромета, а нижний уровень иерархии представлен метеостанциями.

Для повышения надёжности системы связи применяются дублирующие каналы связи. Поэтому каждая линия представляет собой совокупность каналов связи, соединяющих непосредственно двух абонентов.

Основной поток данных — результаты метеорологических измерений, идёт от метеостанций к верхним уровням. Кроме этого, в системе передаётся другая информация — служебная и личная переписка.

Рассмотрим сбор данных гидрометеорологических измерений, поскольку остальные данные, передаваемые в системе, не имеют столь жёстких требований по времени доставки и надёжности.

Особенностью систем сбора гидрометеорологической информации является строгая привязка ко времени и заданное время актуальности данных [3].

На рис. 2 показан порядок сбора данных с метеостанции: T — период сбора данных, Δt — время, в течение которого, данные считаются актуальными, т. е. должны быть доставлены из метеостанции в ЦСД.

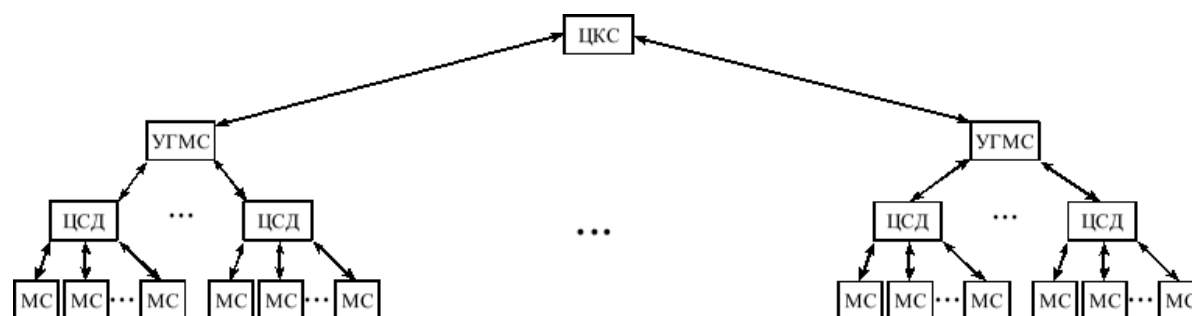


Рис. 1. Структура системы сбора данных Росгидромета

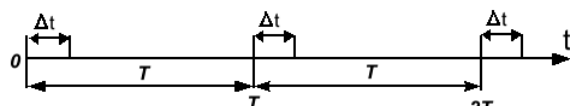


Рис. 2. Привязка сбора данных с метеостанции к времени

Беспроводные каналы связи имеют ряд существенных особенностей по сравнению с проводными или оптоволоконными, среди которых следует выделить следующее:

1. Существенное влияние внешних условий на количество ошибок связи — атмосферы, промышленных и природных помех, солнечной активности и так далее. Влияние указанных внешних условий обычно не поддаётся расчёту и их можно считать случайными возмущениями канала связи. Наиболее целесообразным способом определения количества ошибок в канале связи представляется сбор статистической информации о качестве связи каждого конкретного канала.
2. Влияние момента времени связи. Например, для спутниковых каналов связи — в определённые моменты времени спутники отсутствуют, связь невозможна.
3. Существенные временные затраты на обслуживание канала — установление соединения, переключения приём-передача и т. п.

В рассматриваемых системах сбора гидрометеорологической информации существуют следующие правила обмена информацией между абонентами:

- Сообщение разбивается на пакеты, не превышающие заданный размер.
- Все пакеты сообщения передаются по одному и тому же каналу, выбранному абонентом, передающим сообщение.
- Подтверждения о получении каждого пакета передаются по тому же каналу, что и сообщение.

Оптимизация системы осуществляется по одному из заданных критериев — времени доставки сообщения, стоимости доставки сообщения или по надёжности. Обозначим K_{ij} — совокупность каналов связи, связывающих i -ого и j -ого абонентов. Тогда критерии оптимизации примут следующий вид:

- Оптимизация по времени доставки сообщения предполагает скорейшую доставку сообщений или передачу максимума информации за единицу времени для данной совокупности каналов связи: $t_c = \min(t_c(K_{ij}))$.
- Оптимизация по стоимости доставки сообщения. Предполагает, что стоимость передачи сообщения будет минимально возможной для данной совокупности каналов связи: $C_c = \min(C_c(K_{ij}))$.
- Оптимизация по надёжности. Предполагает, что вероятность ошибки сведётся к минимуму: $P_{errC} = \min(P_{errC}(K_{ij}))$.

Для выбора канала связи в каждый момент времени, необходимо иметь математическое описание системы сбора гидрометеорологической информа-

ции, учитывающее все существенные особенности применяемых каналов связи.

Построение математического описания требует выбора совокупности параметров, описывающих канал связи. Выбранные параметры должны быть заданы как проектные характеристики канала связи, либо измеримы в процессе работы системы. Кроме того, желательно, чтобы число параметров, описывающих канал связи, было минимальным.

Исходя из вышеизложенного, составлен перечень параметров, которые необходимы для описания беспроводного канала связи:

L_p — длина пакета данных.

L_c — длина сообщения.

V_0 — скорость передачи данных при отсутствии ошибок.

C_{i0} — стоимость передачи единицы информации (при тарификации по объёму информации).

C_{t0} — стоимость единицы времени связи (при тарификации по времени связи).

T_{ack} — время ожидания подтверждения.

P_{err0} — вероятность ошибки при передаче единицы информации.

P_d — вероятность доставки сообщения. $0 < P_d < 1$. P_d задаётся исходя из требований надёжности связи.

Особую проблему представляет собой измерение параметра P_{err0} . Поскольку данные передаются пакетами, то ошибка фиксируется при передаче всего пакета. Но практически невозможно определить — была ли это единичная ошибка или несколько ошибок. Для оценки вероятности ошибки при передаче единицы информации P_{err0} введём следующие предположения:

Предположение 1. Ошибкой передачи пакета считается отсутствие подтверждения на данный пакет от принимающего абонента.

Предположение 2. Вероятность ошибки при передаче информации не зависит от положения единицы информации в пакете данных.

Предположение 3. В пакете данных присутствует не более одной ошибки. Если пакет не принят — произошла единичная ошибка. В общем случае данное предположение неверно, но если величину P_{err0} понимать как характеристику вероятности ошибки, то данное предположение допустимо.

Предположение 4. Количество ошибок пропорционально длине передаваемых данных.

Исходя из предположений 1–4, запишем $P_{err} = N_{err}/N$, где N — общее количество переданных пакетов данных, а N_{err} — количество пакетов данных, которые были переданы с ошибками. Обозначим L — суммарную длину переданных сообщений. Поскольку, обычно, в канале связи пакеты имеют сравнимую (а часто и равную) длину, то вероятность ошибки при передаче единицы информации оценим как $P_{err0} = P_{err}/L_{av}$, где $L_{av} = L/N$ — средняя длина пакета. Исходя из этих положений получим:

$$P_{err0} = \frac{N_{err}}{L}. \quad (1)$$

Таким образом, для измерения P_{err0} абонент должен при передаче каждого пакета накапливать статистическую информацию для каждого канала связи – количество пакетов данных, переданных с ошибками:

$$N_{err} = \sum_{i=1 \dots N(t)} N_i;$$

$$\text{где } N_i = \begin{cases} 1, & \text{если } i\text{-й пакет не доставлен} \\ 0, & \text{если } i\text{-й пакет доставлен} \end{cases} \quad (2)$$

и суммарную длину всех пакетов данных:

$$L = \sum_{i=1 \dots N(t)} L_i; \text{ где } L_i - \text{длина } i\text{-го пакета данных.} \quad (3)$$

В формулах (2) и (3) величина $N(t)$ – количество пакетов, переданных к моменту времени t .

Так же необходимо задать минимальное число пакетов $N_{min_{err}}$, прошедших через канал, при превышении которого начинается учёт P_{err0} , то есть необходимо уточнить формулу (1):

$$P_{err0} = \begin{cases} 0, & \text{если } N(t) < N_{min_{err}} \\ \frac{L_{err}}{L}, & \text{если } N(t) \geq N_{min_{err}} \end{cases}. \quad (4)$$

Ограничение (4) необходимо вводить из-за того, что при малом количестве пакетов данных, переданных через канал связи, расчёт вероятности ошибки P_{err0} информации даёт результат далёкий от реального.

В дальнейших рассуждениях положим, что условие $N(t) \geq N_{min}$ выполняется.

Алгоритм оптимизации по времени доставки сообщения:

1. Расчёт для значений $t_C(t)$, где t_C – время доставки сообщения по i -ому каналу связи, рассчитанное на основании выбранных параметров канала связи и накопленной статистической информации, к моменту времени t .
2. Выбирается канал связи, расчётное время доставки по которому наименьшее из всех, и по нему производится обмен данными.

При расчёте $t_C(t)$ учитываются вероятностные параметры канала связи – P_{err0} и P_d .

Положим, что часть пакетов в сообщении передаётся повторно несколько раз из-за ошибок передачи. Количество повторов определяется заданной вероятностью доставки сообщения P_d .

Время доставки сообщения определяется как:

$$t_C = (N_p \cdot T_p + T_{ack}) \cdot R, \quad (5)$$

где t_C – время доставки сообщения, $n_p = L_C/L_p$ – количество пакетов в сообщении, $t_p = (L_p + L_{ack})/V_0$ – время доставки одного пакета данных, R – количество повторов пакетов.

Дальнейшие расчёты ведутся исходя из вероятности ошибки при передаче сообщения. Обоснование расчётов можно найти, например в [4].

Количество повторов пакетов данных рассчитывается исходя из того, что величина ошибки при передаче сообщения не должна превосходить вероятность ошибки доставки сообщения $(1-P_d)$:

$$1 - P_d \geq [1 - (1 - P_{err0})^{(L_p + L_{ack}) \cdot n_p}]^R. \quad (6)$$

Из (6) находим минимальное количество повторов пакетов R при заданной вероятности доставки сообщения P_d :

$$R = \frac{\ln(1 - P_d)}{\ln[1 - (1 - P_{err0})^{(L_p + L_{ack}) \cdot n_p}]}. \quad (7)$$

Подставляя (7) в (5), получаем окончательную формулу (8) расчёта времени доставки пакета с учётом накопленной статистической информации о вероятности ошибок в канале связи:

$$t_C(t) = \left(\frac{L_C}{L_p} \cdot \frac{L_p + L_{ack}}{V_0} + T_{ack} \right) \times \frac{\ln(1 - P_d)}{\ln[1 - (1 - P_{err0})^{(L_p + L_{ack}) \cdot \frac{L_C}{L_p}}]}, \quad (8)$$

где V_0 , L_p , L_C , L_{ack} , T_{ack} , P_d , P_{err0} – параметры, описывающие канал связи и передаваемое сообщение в текущий момент времени t .

Алгоритм оптимизации по стоимости доставки сообщения:

1. Расчёт для значений $C_{Ci}(t)$, где C_{Ci} – стоимость доставки сообщения по i -ому каналу связи.
2. Выбирается канал связи, расчётная стоимость доставки по которому наименьшая из всех и по нему производится обмен данными.

Расчёт C_{Ci} осуществляется на основании тех же предположений, что и расчёт времени доставки. В зависимости от тарификации канала связи – по времени, либо по объёму переданной информации – способы расчёта различны.

Если тарификация производится по времени связи и стоимость единицы времени связи задана как C_0 :

$$C_{C_{i0}}(t) = C_{i0} \cdot t_C(t). \quad (9)$$

Подставляя в (9) время передачи сообщения из (8), получаем формулу для расчёта стоимости передачи сообщения при тарификации по времени связи (10):

$$C_{C_{i0}}(t) = C_{i0} \cdot \left(\frac{L_C}{L_p} \cdot \frac{L_p + L_{ack}}{V_0} + T_{ack} \right) \times \frac{\ln(1 - P_d)}{\ln[1 - (1 - P_{err0})^{(L_p + L_{ack}) \cdot \frac{L_C}{L_p}}]}. \quad (10)$$

При тарификации по объёму занимаемой информации стоимость передачи сообщения:

$$C_{C_{b0}} = (L_C + L_{ack}) \cdot R \cdot C_{b0}. \quad (11)$$

Подставляя в (11) количество повторов пакетов из (7), получаем формулу для расчёта стоимости передачи сообщения при тарификации по объёму переданной информации (12):

$$C_{C_{b0}}(t) = (L_c + L_{ack}) \times \frac{\ln(1 - P_d)}{\ln \left[1 - (1 - P_{err0})^{\frac{(L_p + L_{ack}) \cdot L_c}{L_p}} \right]} \cdot C_{b0}. \quad (12)$$

Оптимизация по надёжности доставки сообщения производится одним из двух способов:

1. Из всей совокупности каналов связи $K_{\text{м}}$ выбирается тот канал, для которого вероятность ошибки при передаче единицы информации P_{err0} является минимальной. Достоинства данного способа — наименьшая стоимость передачи сообщения при максимально возможной надёжности.

СПИСОК ЛИТЕРАТУРЫ

1. Бусленко Н.П. Моделирование сложных систем. — М.: Наука, 1978. — 400с.
2. Сонькин М.А., Слядников Е.Е. Архитектура и общая технология функционирования территориально распределенных аппаратно программных комплексов с пакетной передачей данных // Известия Томского политехнического университета. — 2006. — Т. 309. — № 5. — С. 161–166.

2. Сообщение передают по всей совокупности каналов $K_{\text{м}}$ одновременно. Данный способ позволяет достигнуть большей надёжности доставки сообщения, чем первый способ, но и его стоимость выше.

Предложенный подход к автоматическому выбору канала связи реализован и внедрён в ряд систем по сбору гидрометеорологической информации с низовой сети Росгидромета и Гидрометеослужбы республики Узбекистан.

Выводы

1. Построена математическая модель канала связи с учётом особенностей беспроводных каналов.
2. Модель позволяет оценить вероятность ошибки передачи единицы информации на основе сбора статистических данных в процессе связи и рассчитать время и стоимость передачи сообщения при различных способах тарификации связи.

3. Шамин А.А., Сонькин Д.М. Оценка пропускной способности систем пакетной передачи данных с пульсирующей загрузкой канала связи // Вычислительные технологии. — 2007. — Т. 12. — Спец. вып. № 1. — С. 29–33.
4. Магазинников Л.И. Курс лекций по теории вероятностей. — Томск: Изд-во ТГУ, 1989. — 212 с.